



DIRO
IFT 6150

TRAITEMENT D'IMAGES

RECONNAISSANCE DE FORMES

Max Mignotte

Département d'Informatique et de Recherche Opérationnelle.

Http : [//www.iro.umontreal.ca/~mignotte/ift6150](http://www.iro.umontreal.ca/~mignotte/ift6150)

E-mail : mignotte@iro.umontreal.ca

RECONNAISSANCE DE FORMES

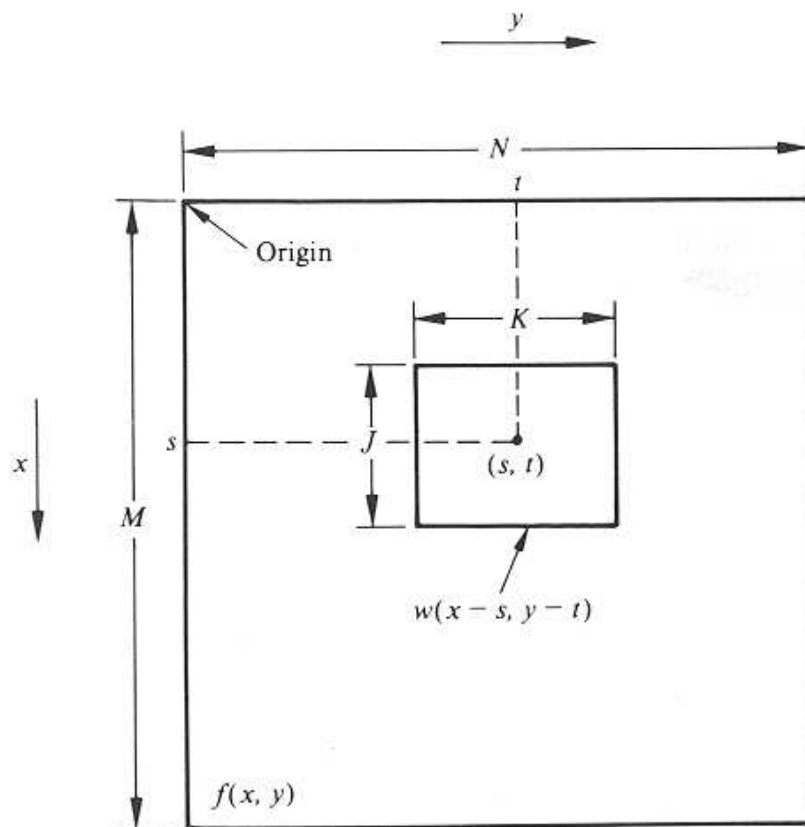
SOMMAIRE

Détection par Corrélacion	2
Principe Général	5
Descripteurs de Fourier	6
Attributs descriptifs de formes	8
Algorithme de clustering non supervisé	10
Algorithme de clustering supervisé	11
Analyse Factorielle Discriminante	12

RECONNAISSANCE DE FORMES

DÉTECTION PAR CORRÉLATION (1)

- Le but de la détection d'objets est de localiser dans une image $f(x, y)$ de dimension $M \times N$ des objets ou formes représentés par une sous image $w(x, y)$ de dimension $J \times K$ ($J \leq M$ et $K \leq N$)
- L'approche par corrélation consiste à déplacer l'objet $w(x, y)$ sur l'image $f(x, y)$ et de calculer une fonction de ressemblance entre $w(x, y)$ et $f(x, y)$ au voisinage de chaque point



$$C(s, t) = \sum_{x=s-J/2}^{s+J/2} \sum_{y=t-K/2}^{t+K/2} f(x, y) w(x - s, y - t) = f(x, y) * w(-x, -y)$$

- Sensible aux changements d'amplitude de w et f (contraste et brillance)

RECONNAISSANCE DE FORMES

DÉTECTION PAR CORRÉLATION (2)

► Mesure de corrélation insensible aux changements de contraste et de brillance

$$\rho(s, t) = \frac{\sum_{x=s-J/2}^{s+J/2} \sum_{y=t-K/2}^{t+K/2} [f(x, y) - \bar{f}(x, y)] [w(x-s, y-t) - \bar{w}]}{\sqrt{AB}}$$

avec

$$A = \sum_{x=s-J/2}^{s+J/2} \sum_{y=t-K/2}^{t+K/2} [f(x, y) - \bar{f}(x, y)]^2$$

et

$$B = \sum_{x=s-J/2}^{s+J/2} \sum_{y=t-K/2}^{t+K/2} [w(x-s, y-t) - \bar{w}]^2$$

Exemple



(a)



(b)



(c)

RECONNAISSANCE DE FORMES

DÉTECTION PAR CORRÉLATION (3)

- ▶ Onéreuse en temps machine
- ▶ Solution : accomplir la corrélation dans le domaine spectrale

$$C(s, t) = f(x, y) * w(-x, -y)$$

$$\begin{array}{c} \nearrow \\ \mathbf{F}^{-1} \left| \right. \\ \searrow \end{array} \mathbf{F}$$

$$C(u, \nu) = F(u, \nu) \cdot W^*(u, \nu)$$

Application

- ▶ Détection & suivie de formes dans des séquences d'images



frame 4



frame 10



frame 16



frame 24

RECONNAISSANCE DE FORMES

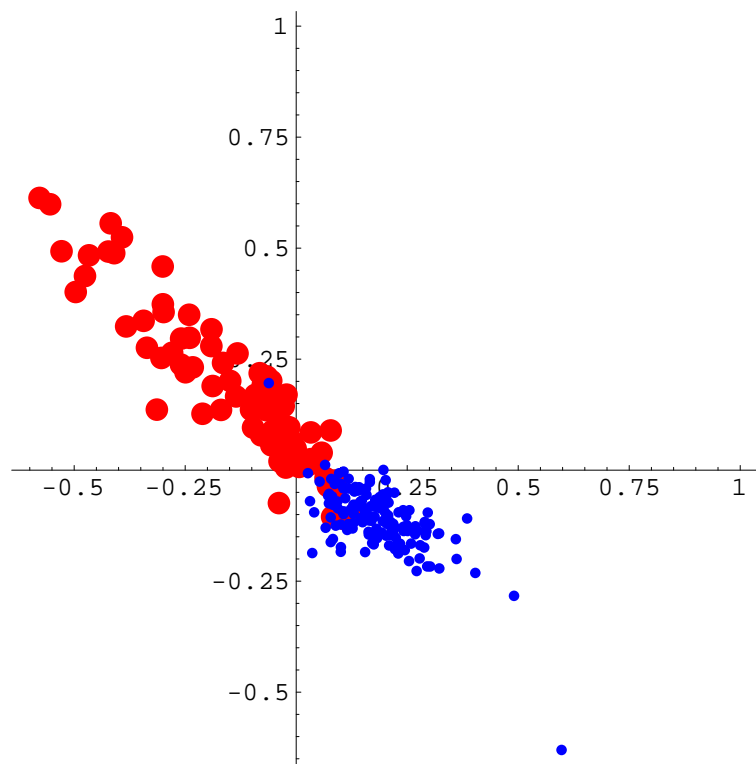
PRINCIPE GÉNÉRAL

1. Reconnaissance de forme par corrélation (Matching)

► Cf. slides précédents

2. Méthodes multidimensionnelles

► Les méthodes multidimensionnelles consistent à associer à chaque forme (ou texture) un vecteur de q attributs et à rechercher dans cette espace de représentation (de dimension q) le cluster ou la classe à laquelle il appartient



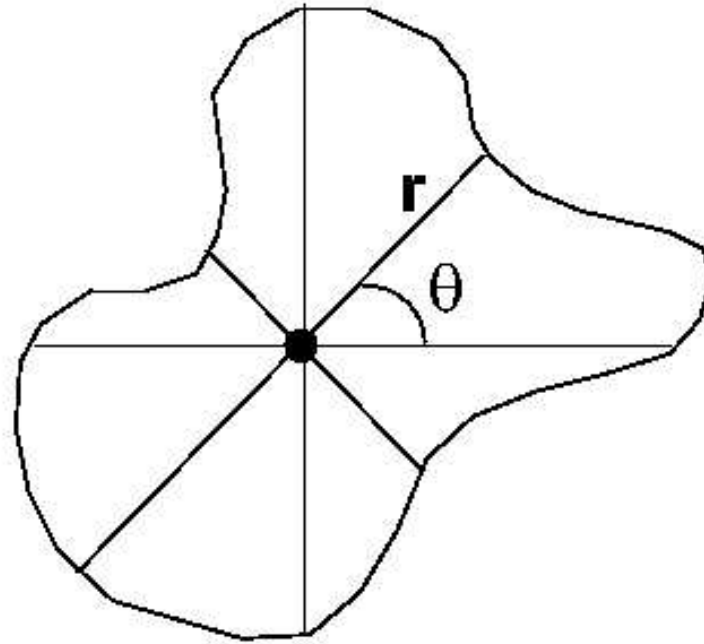
Reconnaissance de Texture

► Cf. Chap. Texture & Traitement d'images couleurs

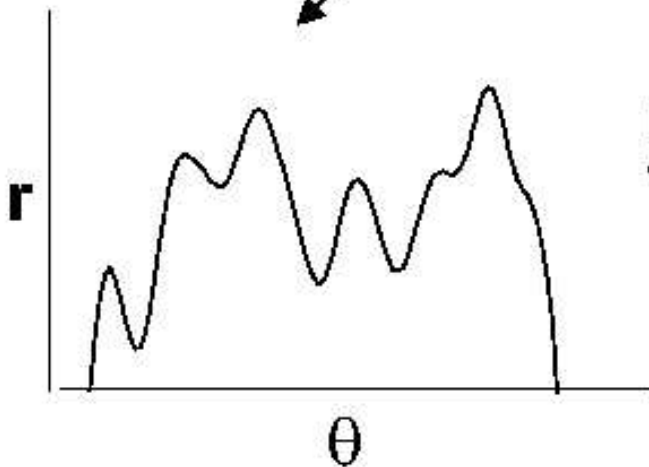
RECONNAISSANCE DE FORMES

DESCRIPTEURS DE FOURIER (1)

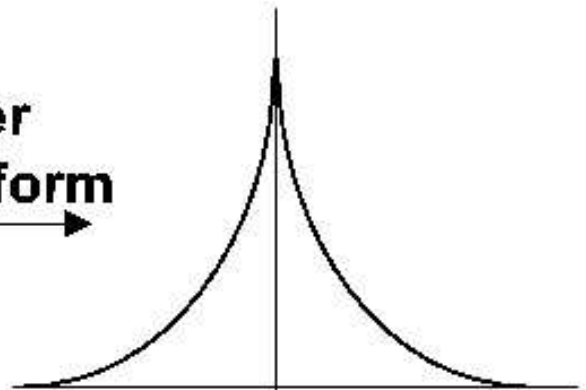
Descripteurs de Fourier



**Boundary
Representation**



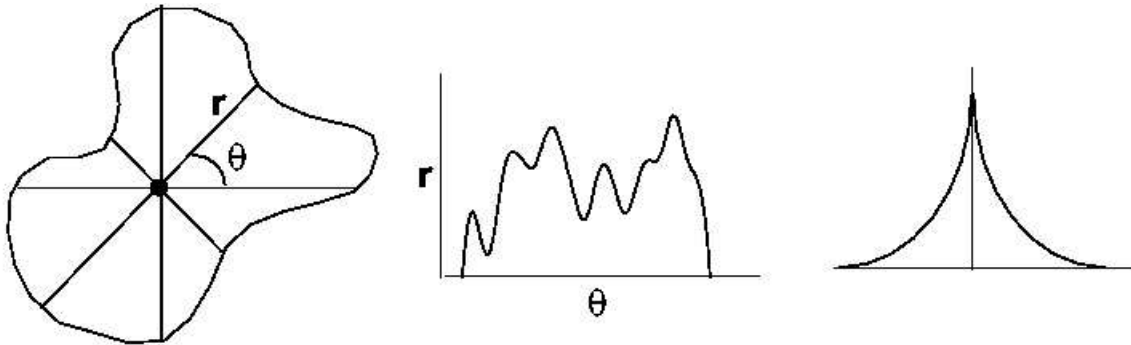
**Fourier
Transform**



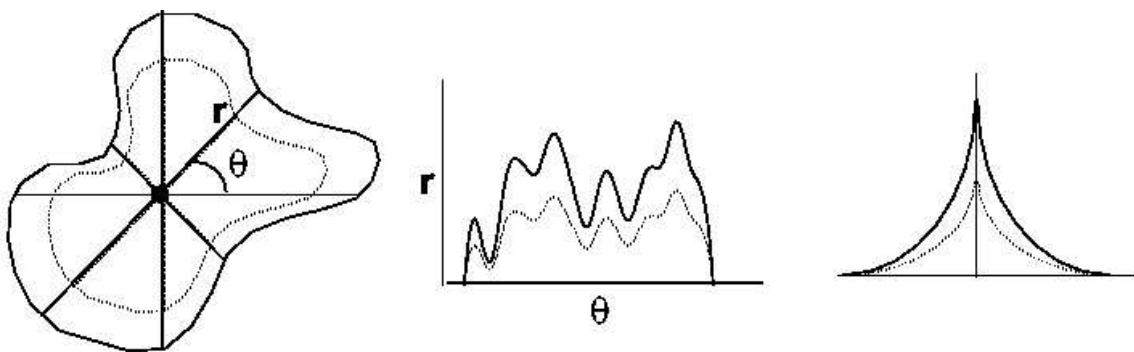
RECONNAISSANCE DE FORMES

DESCRIPTEURS DE FOURIER (2)

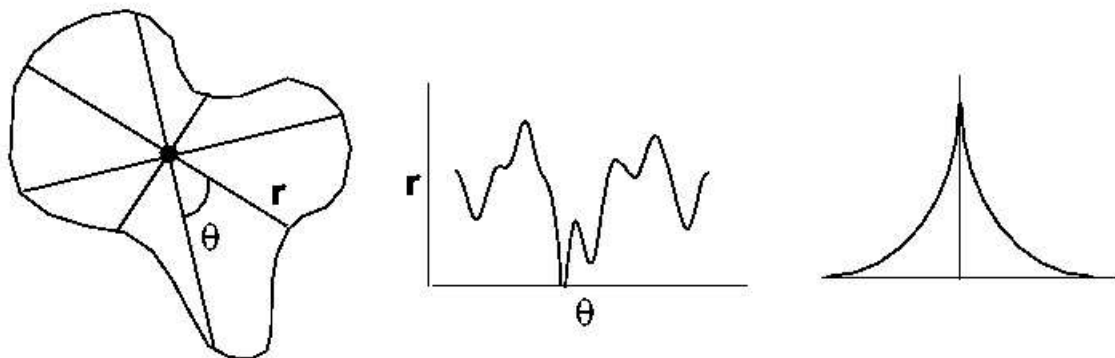
Translation



Facteur d'échelle



Rotation



RECONNAISSANCE DE FORMES

ATTRIBUTS DESCRIPTIFS DE FORMES (3)

Compacité

- Mesure le degré de compacité d'une forme

$$C = \frac{4\pi S}{l^2}$$

où l et S désignent le périmètre et l'aire de la surface délimité par le contour (= 1 pour un cercle)

Élongation

- Paramètre donné par les valeurs propres de la matrice de corrélation (cf. Transformée de Hotteling)

$$C = \begin{pmatrix} c_{xx} & c_{xy} \\ c_{xy} & c_{yy} \end{pmatrix}$$

On obtient

$$\begin{aligned} \sigma_1 &= \frac{1}{2} \left(c_{xx} + c_{yy} + \sqrt{(c_{xx} - c_{yy})^2 + 4c_{xy}^2} \right) \\ \sigma_2 &= \frac{1}{2} \left(c_{xx} + c_{yy} - \sqrt{(c_{xx} - c_{yy})^2 + 4c_{xy}^2} \right) \end{aligned}$$

Orientation

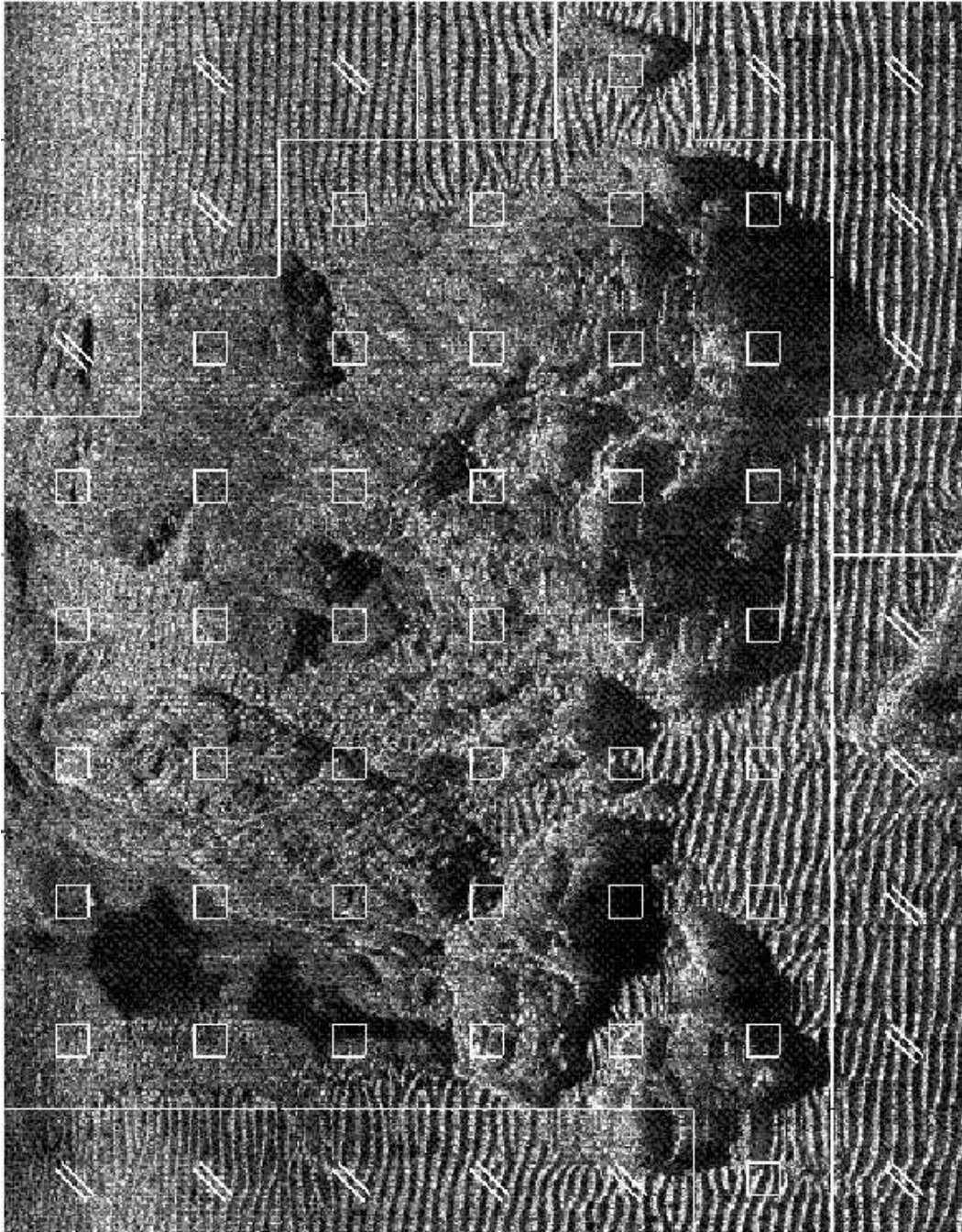
- Paramètre calculé à partir des valeurs de la matrice de corrélation (cf. Transformée de Hotteling)

$$\alpha = \arctan \left(\frac{\sigma_1 - c_{xx}}{c_{xy}} \right)$$

RECONNAISSANCE DE FORMES

ATTRIBUTS DESCRIPTIFS DE FORMES (4)

Exemple : Classification des fonds marins



RECONNAISSANCE DE FORMES

ALGORITHME DE CLUSTERING NON SUPERVISÉ

Algorithme K-Moy

1. On choisit les K premiers échantillons comme étant les K centres $(c_1^{[1]}, \dots, c_K^{[1]})$ des K groupements

$$c_i^{[1]} = \mathbf{x}_i, \quad 1 \leq i \leq K$$

2. A l'itération p , on associe l'échantillon \mathbf{x}_l , ($1 \leq l \leq M$) au groupement C_i (noté $C_i^{[p]}$) si

$$\| \mathbf{x}_l - c_i^{[p]} \| < \| \mathbf{x}_l - c_j^{[p]} \|, \quad \forall j \neq i$$

(On associe l'échantillon \mathbf{x}_l au groupement dont le centre lui est le plus proche)

3. $C_i^{[p]}$ correspond au groupement i constitué à l'itération p après l'étape 2. On détermine le nouveau centre de chaque groupement par

$$c_i^{[p+1]} = \frac{1}{N_i} \sum_{\mathbf{x} \in C_i^{[p]}} \mathbf{x}, \quad \forall i$$

avec N_i le nombre d'échantillons de $C_i^{[p]}$

4. Si les centres des groupements à l'itération p sont les mêmes que ceux obtenus à l'itération précédente, alors l'algorithme a convergé si $c_i^{[p+1]} = c_i^{[p]}, \quad \forall i$, fin de l'algorithme sinon on retourne en 2.

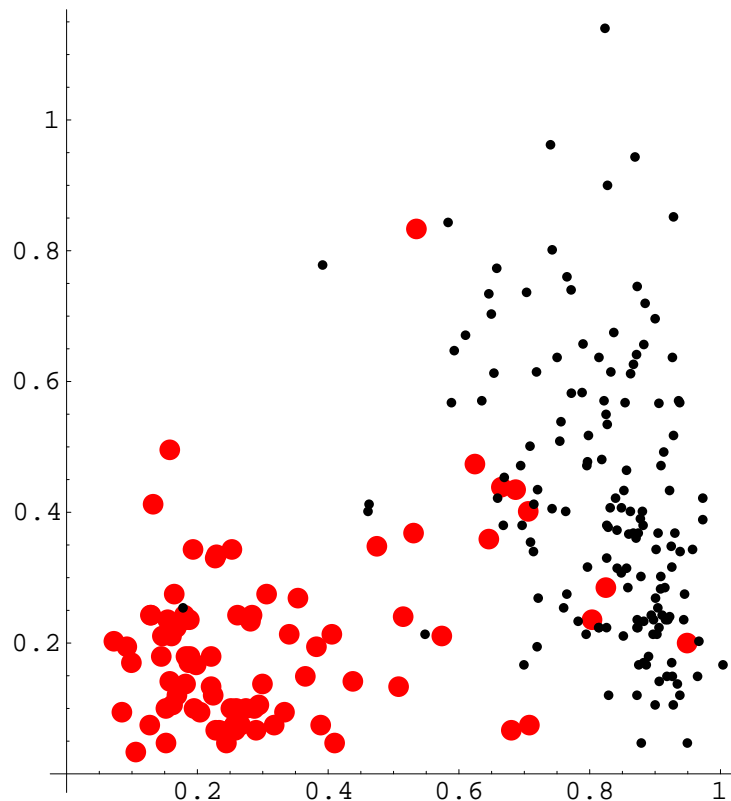
RECONNAISSANCE DE FORMES

ALGORITHME DE CLUSTERING SUPERVISÉ

- ▶ Supposent la connaissance *a priori* de l'appartenance de chaque échantillon de l'ensemble d'apprentissage à une classe donnée

Algorithme K-plus proches voisins

1. Pour un K donné, on affecte l'échantillon x_i , la classe majoritaire parmi les K plus proches voisins
- ▶ Si K est faible, le vote risque d'être perturbé par le bruit
 - ▶ Coût de stockage important (les éléments de l'ensemble d'apprentissage doivent être stockés)
 - ▶ Coût élevé de la recherche des Kppv ($O(K^2n \log n)$)



RECONNAISSANCE DE FORMES

ANALYSE FACTORIELLE DISCRIMINANTE (1)

► L'AFD a pour but de trouver les axes factoriels qui discriminent au mieux les K classes de l'ensemble d'apprentissage (i.e., les projections des échantillons d'une même classe soient les + concentrées possible)

-
- L'ensemble d'apprentissage est stocké dans une matrice X de M lignes (représentant les échantillons) et de q colonnes (représentant les attributs)
 - Échantillon : $\mathbf{x}_i = (x_i^j)_{1 \leq j \leq q}$, vecteur de ses q attributs
 - Valeurs du j ième attribut, stockées dans $\mathbf{x}^j = (x_i^j)_{1 \leq i \leq M}$
 - Chaque individu appartient à une classe \mathcal{C}_k , $k = 1, \dots, K$

1. La moyenne de l'attribut j

$$\bar{x}^j = \sum_{i=1}^M x_i^j$$

2. $\bar{\mathbf{x}}$ = Le vecteur des moyennes de chaque attributs

3. $\bar{\mathbf{x}}_k$ = Le vecteur des moyennes de chaque attributs des échantillons de \mathcal{C}_k

4. N_k = Le cardinal de \mathcal{C}_k

RECONNAISSANCE DE FORMES

ANALYSE FACTORIELLE DISCRIMINANTE (2)

Soit \mathbf{T} , \mathbf{B} et \mathbf{W} les matrices de variances totale, inter-classe et intraclasse,

Matrice de Variance totale

$$\mathbf{T} = \frac{1}{M} \sum_{i=1}^M (\mathbf{x}_i - \bar{\mathbf{x}}) (\mathbf{x}_i - \bar{\mathbf{x}})^t$$

Matrice de Variance intraclasse

$$\mathbf{W} = \frac{1}{M} \sum_{k=1}^K \sum_{m \in \mathcal{C}_k} (\mathbf{x}_m - \bar{\mathbf{x}}_k) (\mathbf{x}_m - \bar{\mathbf{x}}_k)^t$$

Matrice de Variance interclasse

$$\mathbf{B} = \frac{1}{M} \sum_{k=1}^K N_k (\bar{\mathbf{x}}_k - \bar{\mathbf{x}}) (\bar{\mathbf{x}}_k - \bar{\mathbf{x}})^t$$

Ces trois matrices sont liées par la relation

$$\mathbf{T} = \mathbf{W} + \mathbf{B}$$

► Le problème est de déterminer le système orthonormé de vecteurs \mathbf{u}_i $i = 1, \dots, r$ de \mathbb{R}^q , ($r < q$) pour lequel l'inertie intraclasse est minimale et l'inertie interclasse maximale

On cherche les vecteurs \mathbf{u}_i qui minimisent

$$\frac{\mathbf{u}_i^t \mathbf{W} \mathbf{u}_i}{\mathbf{u}_i^t \mathbf{T} \mathbf{u}_i}$$

sous la contrainte $\mathbf{u}_i^t \mathbf{T} \mathbf{u}_i = 1$

RECONNAISSANCE DE FORMES

ANALYSE FACTORIELLE DISCRIMINANTE (3)

comme $\mathbf{u}_i^t \mathbf{T} \mathbf{u}_i = \mathbf{u}_i^t \mathbf{B} \mathbf{u}_i + \mathbf{u}_i^t \mathbf{W} \mathbf{u}_i$

Ceci est équivalent à déterminer les \mathbf{u}_i qui maximisent $\mathbf{u}_i^t \mathbf{B} \mathbf{u}_i$ sous la contrainte $\mathbf{u}_i^t \mathbf{T} \mathbf{u}_i = 1$

Les axes factoriels discriminants sont donnés par les vecteurs \mathbf{v}_i vérifiant

$$\mathbf{B} \mathbf{T}^{-1} \mathbf{u}_i = \lambda_i \mathbf{u}_i$$

Ces vecteurs sont donc les vecteurs propres de la matrice $\mathbf{B} \mathbf{T}^{-1}$. On se limite généralement aux r (r fixé) plus grande valeurs propres telles qu'elles constitue une proportion importante de la trace de la matrice $\mathbf{B} \mathbf{T}^{-1}$

$$\text{Tr}(\mathbf{B} \mathbf{T}^{-1}) = \sum_i^r \lambda_i$$

- Il est donc possible de diminuer la taille de l'espace discriminant par cette méthode
- Dans ce sous-espace, la moyenne des variances de classes est minimale et les moyennes de classes sont les plus dispersées possible.